

Czynniki determinujące wydatki gospodarstw domowych emerytów na rekreację i kulturę

Uczestnictwo w rekreacji oraz kulturze można określić za pomocą wielkości wydatków przeznaczonych przez gospodarstwa domowe na ten cel. Dane statystyczne o przeciętnych miesięcznych wydatkach przypadających na osobę w gospodarstwie domowym można uzyskać m.in. z Badań Budżetów Gospodarstw Domowych (BGD). W badaniu wyróżnia się 46 rodzajów wydatków na rekreację i kulturę. Wśród nich znajdują się m.in. wydatki na sprzęt służący rekreacji i kulturze, sprzęt turystyczny oraz na turystykę zorganizowaną.

Celem artykułu jest identyfikacja czynników wpływających na podejmowanie decyzji o wielkości wydatków na rekreację i kulturę w gospodarstwach domowych emerytów. Podstawę informacyjną badań stanowiły informacje liczbowe o indywidualnych gospodarstwach domowych pochodzące z badania BGD przeprowadzonego w 2009 r. przez GUS. Dane te zostały odpowiednio przetworzone i skojarzone z wynikami badania ankietowego pt. *Turystyka i wypoczynek w gospodarstwach domowych*, przeprowadzonego przez GUS w tym samym roku. Zintegrowany zbiór danych obejmował 1308 gospodarstw domowych emerytów.

W badaniu wzięto pod uwagę głównie czynniki kategoryzacyjne i z tego względu do wyboru optymalnego zbioru czynników charakteryzujących wydatki gospodarstw domowych na rekreację i kulturę wykorzystano analizę log-liniową. Pozwala ona bardziej precyzyjnie opisać zależności między zmiennymi kategoryzacyjnymi w porównaniu z miernikami stosowanymi do oceny współzależności cech jakościowych. Dodatkowym atutem analizy log-liniowej jest możliwość uwzględnienia wpływu interakcji pomiędzy zmiennymi.

METODA

Analiza log-liniowa znajduje zastosowanie szczególnie w naukach rolniczych (Szwedziak, 2005), biologicznych i medycznych, w których wiele zjawisk ma charakter jakościowy (Żołnierczuk-Kieliszek i in., 2006). Wykorzystywana jest również w badaniach ekonomicznych, np. do wyboru czynników opisujących sytuację ekonomiczną gospodarstw domowych (Salamaga, 2008) czy też do estymacji modeli cen oraz indeksów cen nieruchomości mieszkaniowych (Tomczyk, Widłak, 2010).

Punktem wyjścia w analizie log-liniowej jest liczebność oczekiwana dla poszczególnych poziomów zmiennych kategoryzacyjnych (Dobosz, 2004). Jeżeli

człony interakcyjne są nieistotne, a liczebność w klasach zależy tylko od czynników głównych, oznacza to, że brak jest wzajemnych oddziaływań pomiędzy zmiennymi kategoryzacyjnymi.

Dopasowywany model log-liniowy ma charakter hierarchiczny, a zatem jeżeli określony człon interakcyjny jest włączony do modelu, to wszystkie pozostałe kombinacje czynników występujących w tym członie muszą być uwzględnione w modelu.

W modelu log-liniowym przyjmuje się, że logarytm naturalny wartości oczekiwanej liczebności w komórce w tabeli niezależności jest liniową funkcją czynników. Przyjmując, że mamy następujące czynniki: C_1 o p_{c_1} poziomach, C_2 o p_{c_2} poziomach, ..., C_m o p_{c_m} poziomach, uogólniona postać modelu jest następująca (Salamaga, 2008):

$$\ln(\hat{n}^C) = \bar{n} + \sum_{i=1}^{p_{c_1}} \lambda_i^{C_1} + \dots + \sum_{j=1}^{p_{c_2}} \lambda_{ij}^{C_1 C_2} + \dots + \sum_{k=1}^{p_{c_3}} \lambda_{ijk}^{C_1 C_2 C_3} + \dots \quad (1)$$

gdzie:

- $\hat{n}^C$ — jest wektorem $[\hat{n}_i^{C_1}, \hat{n}_{ij}^{C_1 C_2}, \dots]$, przy czym $\hat{n}_i^{C_1}$ oznacza oczekiwaną liczebność i -tego poziomu czynnika C_1 , $\hat{n}_{ij}^{C_1 C_2}$ przewidywaną liczebność dotyczącą interakcji i -tego poziomu czynnika C_1 oraz j -tego poziomu czynnika C_2 itd.,
- $\bar{n}$ — średnia z logarytmów naturalnych obserwowanej liczebności wyznaczona na podstawie wzoru:

$$\bar{n} = \frac{1}{n} \sum_C \sum_i \ln(n_i^C) \quad (2)$$

gdzie:

- $\sum_C$ — sumowanie po wszystkich czynnikach,
- $\sum_i$ — sumowanie po wszystkich poziomach czynników,

- $\lambda_i^{C_1}$ — wskaźnik i -tego poziomu czynnika C_1 ,
- $\lambda_{ij}^{C_1 C_2}$ — wskaźnik interakcji rzędu drugiego i -tego poziomu czynnika C_1 oraz j -tego poziomu czynnika C_2 ,
- $\lambda_{ijk}^{C_1 C_2 C_3}$ — wskaźnik interakcji rzędu trzeciego pomiędzy i -tym poziomem czynnika C_1 , j -tym poziomem czynnika C_2 oraz k -tym poziomem czynnika C_3 itd.

Poprawnie zbudowany model log-liniowy umożliwia najlepszą predykcję liczebności, przy uwzględnieniu w modelu jak najmniejszej liczby interakcji. Uwzględnienie wszystkich czynników i ich interakcji daje model najlepiej dopasowany, lecz nie zawsze wygodny w praktyce, gdyż wpływ niektórych czynników i interakcji może być niewielki w porównaniu z pozostałymi składnikami modelu. Miarą dopasowania modelu do wyników są statystyka χ^2 Pearsona oraz χ^2 największej wiarygodności. Test χ^2 Pearsona polega na porównaniu liczebności zaobserwowanej (częstości) f_{ij} z liczebnością oczekiwaną $n\pi_{ij}$, przy założeniu prawdziwości hipotezy zerowej (o braku związku pomiędzy zmiennymi). Gdyby nie było żadnej zależności pomiędzy zmiennymi, wówczas powinniśmy oczekiwać mniej więcej takiej samej liczebności oczekiwanej i obserwowanej. Test χ^2 staje się istotny, w miarę jak liczebność oczekiwana zaczyna się różnić od liczebności obserwowanej. Miarą różnic pomiędzy liczebnością teoretyczną i zaobserwowaną jest statystyka (Dobosz, 2004):

$$\chi^2 = \sum_{i=1}^k \sum_{j=1}^p \frac{(f_{ij} - n\pi_{ij})^2}{n\pi_{ij}} \quad (3)$$

gdzie k, p — liczby kategorii rozpatrywanych zmiennych.

Ogólnie

$$\chi^2 = \sum_{i_1=1}^{p_1} \sum_{i_2=1}^{p_2} \dots \eta_{i_1 i_2 \dots} \quad (4)$$

gdzie

$$\eta_{i_1 i_2 \dots} = \frac{(f_{i_1 i_2 \dots} - n\pi_{i_1 i_2 \dots})^2}{n\pi_{i_1 i_2 \dots}} \quad (5)$$

przy czym p_j to liczba poziomów zmiennej $j=1, 2, \dots$

Test χ^2 największej wiarygodności sprawdza tę samą hipotezę co test χ^2 Pearsona, ale jego sposób obliczania oparty jest na teorii największej wiarygodności. W podejściu tym wykorzystuje się własności statystyki, jaką jest iloraz wiarygodności. Dla każdej komórki tabeli wielodzielczej o dowolnym wymiarze wyznacza się tzw. składnik największej wiarygodności $c_{(ijk\dots)}$ proporcjonalny do logarytmu naturalnego ilorazu liczebności zaobserwowanej $f_{(ijk\dots)}$ do oczekiwanej (dopasowanej modelem) $\hat{f}_{(ijk\dots)}$ według wzoru:

$$c_{(ijk\dots)} = 2 f_{(ijk\dots)} \ln \frac{f_{(ijk\dots)}}{\hat{f}_{(ijk\dots)}} \quad (6)$$

Statystykę testową największej wiarygodności wyznacza się jako sumę składników najwyższej wiarygodności ze wszystkich komórek tabeli wielodzielczej, czyli jako:

$$\chi^2_{NW} = 2 \sum_i^a \sum_j^b \dots c_{(ijk\dots)} \quad (7)$$

W praktyce statystyka χ^2 największej wiarygodności generuje wyniki zbliżone do statystyki χ^2 Pearsona.

Po ustaleniu rzędu interakcji należy włączyć je do modelu. Pojawia się jednak pytanie, które z interakcji danego rzędu (pomiędzy którymi czynnikami) powinny być uwzględnione w modelu? W rozstrzygnięciu tego dylematu pomocna może okazać się analiza zależności cząstkowych i brzegowych. Zależności cząstkowe informują o tym, czy odpowiednie interakcje są istotne statystycznie, w sytuacji gdy wszystkie inne czynniki tego samego stopnia są już w modelu. Weryfikujemy je testem χ^2 Pearsona. Z kolei analiza zależności brzegowych wskazuje na istnienie wpływu określonych interakcji, w przypadku gdy w modelu jeszcze nie uwzględniono żadnej interakcji tego samego rzędu. Przeprowadza się ją przy użyciu testu wiarygodności χ^2 .

Modele log-liniowe są bardzo podobne do modeli zmiennych ilościowych używanych w analizie wariancji, różnica dotyczy interpretacji. Ponadto w analizie log-liniowej zwraca się uwagę głównie na efekty interakcji, a nie na efekty główne.

WYNIKI BADAŃ

Badaniu poddano 1308 gospodarstw domowych emerytów. Sytuacja ekonomiczna respondentów była zróżnicowana. W zdecydowanej większości gospodarstw (82%) nie było osób pracujących. Tylko niespełna 22% spośród nich miało oszczędności. Zbliżony odsetek gospodarstw korzystał z kredytów i pożyczek. W ponad 32% badanych gospodarstw domowych dochód na jednego członka rodziny nie przekraczał 1 tys. zł. Jednocześnie tylko 11% gospodarstw deklarowało dochód powyżej 2 tys. zł.

Większość badanych gospodarstw emerytów była słabo wyposażona w sprzęt służący zaspokojeniu potrzeb rekreacyjno-kulturowych. Tylko nieco ponad 3% respondentów posiadało domek letniskowy i kamerę wideo. Działka rekreacyjna była własnością ponad 12% gospodarstw, ok. 24% miało komputer z dostępem do Internetu oraz magnetowid lub odtwarzacz. Samochód posiadało ponad 37% badanych. Najwięcej respondentów miało rower (54,1%), urządzenie do odbioru TV satelitarnej lub kablowej (53,13%) oraz aparat fotograficzny (40,4%). W prawie 84% ankietowanych gospodarstw stwierdzono wydatki na rekreację i kulturę.

Największą popularnością wśród członków gospodarstw domowych emerytów cieszyły się następujące sposoby spędzania czasu wolnego: słuchanie radia, muzyki, oglądanie telewizji, wideo (34,2%), czytanie (17,4%), wypoczynek bierny — opalanie się, relaks (14,8%), udział w spotkaniach towarzyskich (5,0%), praca na działce (11,6%), praktyki religijne (5,0%), uprawianie sportu i ćwiczenia fizyczne (4,4%), hobby, pogłębianie wiedzy (2,5%), odwiedzanie muzeów, wystaw i zabytków, uczestnictwo w imprezach objazdowych, zwiedzanie (1,4%), inne (3,7%).

Z uwagi na cel artykułu wyróżniono dychotomiczną zmienną zależną zdefiniowaną jako ponoszenie wydatku na rekreację i kulturę (Y). W zbiorze zmiennych niezależnych uwzględniono następujące cechy gospodarstw domowych emerytów:

- X_1 — pracujący (tak, nie),
- X_2 — liczba osób pobierających emeryturę lub rentę (od 1 do 2 osób, powyżej 2 osób),
- X_3 — dochód w przeliczeniu na jednego członka (do 1 tys. zł, od 1 do 2 tys. zł, powyżej 2 tys. zł),
- X_4 — oszczędności (tak, nie),
- X_5 — zadłużenie (tak, nie),
- X_6 — sytuacja mieszkaniowa respondenta (mieszkanie własne, inna forma prawna użytkowania),
- X_7 — powierzchnia mieszkania (mała — do 40 m², średnia — od 40 m² do 60 m², duża — powyżej 60 m²),
- X_8 — liczba pokoi w mieszkaniu (jeden pokój, dwa pokoje, trzy pokoje, powyżej trzech pokoi);

posiadanie:

- X_9 — aparatu fotograficznego (tak, nie),
- X_{10} — komputera z dostępem do Internetu (tak, nie),
- X_{11} — radia lub radiomagnetofonu (tak, nie),
- X_{12} — urządzenia do odbioru TV satelitarnej lub kablowej (tak, nie),
- X_{13} — działki rekreacyjnej (tak, nie),
- X_{14} — samochodu osobowego (tak, nie),
- X_{15} — przyczepy kempingowej (tak, nie),
- X_{16} — roweru (tak, nie).

Wymienione zmienne przedstawiono w postaci zmiennych kategoryzacyjnych i dlatego (jak już wspomniano) do wyboru optymalnego zbioru czynników charakteryzujących wydatki na rekreację i kulturę posłużono się analizą log-liniową.

W pierwszym kroku zbadano, które ze zmiennych wykazują istotny związek ze zmienną zależną. Do tego celu wykorzystano test niezależności χ^2 . W tabl. 1 podano wartość statystyki testującej χ^2 oraz prawdopodobieństwo testowe p .

TABL. 1. WARTOŚĆ STATYSTYKI χ^2 ORAZ PRAWDOPODOBIEŃSTWO TESTOWE p

Zmienne statystyczne	Wartości statystyki χ^2	Wartości prawdopodobieństwa testowego p
X_1	5,320	0,02106
X_2	2,610	0,10618
X_3	42,320	0,00000
X_4	23,111	0,00000
X_5	13,962	0,00019
X_6	13,330	0,00026
X_7	1,159	0,56014
X_8	11,414	0,00332
X_9	62,210	0,00000
X_{10}	55,860	0,00000
X_{11}	1,170	0,27892
X_{12}	137,720	0,00000
X_{13}	6,940	0,03108
X_{14}	25,950	0,00000
X_{15}	0,660	0,72050
X_{16}	11,740	0,00283

U w a g a. Pogrubioną czcionką zaznaczono zmienne niezależne.

Ź r ó d ł o: obliczenia własne.

Na podstawie przeprowadzonych obliczeń można stwierdzić, że zmienna określająca wystąpienie wydatków na rekreację i kulturę nie jest powiązana z czterema zmiennymi: z liczbą osób pobierających emeryturę lub rentę (X_2), powierzchnią mieszkania (X_7), posiadaniem radia lub radiomagnetofonu (X_{11}) oraz posiadaniem przyczepy kempingowej (X_{15}). Zmienne te pominięto w dalszym badaniu.

W programie *Statistica* można analizować maksymalnie tabele siedmiodzielcze w celu znalezienia czynników wpływających istotnie na podjęcie decyzji o wydatkach na rekreację i kulturę w gospodarstwach domowych emerytów. Dokonano więc podziału analizowanych zmiennych na dwie grupy. W pierwszej znalazły się zmienne charakteryzujące sytuację ekonomiczną gospodarstwa domowego, w drugiej — dotyczące wyposażenia gospodarstw w sprzęt służący zaspokojeniu potrzeb rekreacyjnych i kulturowych.

Do pierwszego modelu log-liniowego wprowadzono następujące zmienne: X_1 , X_3 , X_4 , X_5 , X_6 , X_8 . W celu dokonania specyfikacji modelu należy określić rząd interakcji zmiennych. Wykonano zatem testowanie wszystkich interakcji. Wyniki zamieszczono w tabl. 2. Pokazują one, że w modelu log-liniowym oprócz czynników głównych powinno się uwzględnić interakcje co najwyżej drugiego rzędu. Wskazują na to wartości prawdopodobieństw testowych p odpowiadające wartościom statystyki testowej χ^2 największej wiarygodności i χ^2 Pearsona.

Aby ocenić, które z interakcji powinny być włączone do modelu log-liniowego wykorzystano testy cząstkowe i brzegowe (tabl. 3). Z uwagi na znaczną liczbę możliwych interakcji w przypadku siedmiu czynników, ograniczono się do przedstawienia interakcji maksymalnie rzędu drugiego, ponieważ żadna z interakcji wyższych rzędów nie okazała się statystycznie istotna w sensie testów cząstkowego i brzegowego.

TABL. 2. WYNIKI TESTÓW INTERAKCJI MIĘDZY ZMIENNYMI $Y, X_1, X_3, X_4, X_5, X_6, X_8$

Stopień interakcji	Stopnie swobody	Wartość χ^2 największej wiarygodności		Wartość χ^2 Pearsona	
			prawdopodobień- stwo p		prawdopodobień- stwo p
1	9	2555,601	0,00000	7046,597	0,00000
2	34	489,920	0,00000	610,460	0,00000
3	70	74,777	0,32601	81,696	0,16014
4	85	50,199	0,99904	48,125	0,99957
5	61	17,167	1,00000	17,039	1,00000
6	24	8,448	0,99856	8,530	0,99844
7	4	1,129	0,88957	1,140	0,88789

U w a g a. Jak przy tabl. 1.

Ź r ó d ł o: jak przy tabl. 1.

**TABL. 3. NIEKTÓRE WYNIKI TESTÓW CZĄSTKOWYCH I BRZEGOWYCH
MIĘDZY ZMIENNYMI $Y, X_1, X_3, X_4, X_5, X_6, X_8$**

Czynniki	Stopnie swobody	Związek cząstkowy χ^2	Wartość p w związku cząstkowym	Związek brzegowy χ^2	Wartość p w związku brzegowym
Y	1	577,652	0,000	577,652	0,000
X_1	1	483,873	0,000	483,873	0,000
X_3	2	378,227	0,000	378,227	0,000
X_4	1	388,726	0,000	388,726	0,000
X_5	1	371,147	0,000	371,147	0,000
X_6	1	175,750	0,000	175,750	0,000
X_8	2	302,311	0,000	302,311	0,000
YX_1	1	0,218	0,640	0,066	0,797
YX_3	2	16,717	0,000	23,671	0,000
YX_4	1	1,711	0,191	4,940	0,026
YX_5	1	0,581	0,446	1,825	0,177
YX_6	1	3,416	0,065	4,039	0,044
YX_8	2	19,998	0,000	19,257	0,000
X_1X_3	2	35,188	0,000	33,567	0,000
X_1X_4	1	13,542	0,000	17,447	0,000
X_1X_5	1	14,743	0,000	21,176	0,000
X_1X_6	1	0,066	0,798	0,387	0,534
X_1X_8	2	16,983	0,000	18,314	0,000
X_3X_4	2	23,833	0,000	36,037	0,000
X_3X_5	2	8,915	0,012	16,922	0,000
X_3X_6	2	19,841	0,000	25,559	0,000
X_3X_8	4	6,795	0,147	6,746	0,150
X_4X_5	1	34,539	0,000	48,464	0,000
X_4X_6	1	9,984	0,002	16,060	0,000
X_4X_8	2	0,606	0,738	0,096	0,953
X_5X_6	1	0,072	0,788	1,540	0,215
X_5X_8	2	7,305	0,026	8,299	0,016
X_6X_8	2	29,903	0,000	25,944	0,000

U w a g a. Pogrubioną czcionką zaznaczono interakcje, dla których zależność cząstkowa i brzegowa jest istotna.

Ź r ó d ł o: jak przy tabl. 1.

Rezultaty testów cząstkowych i brzegowych wskazują na potrzebę uwzględnienia w modelu log-liniowym siedmiu czynników głównych oraz tych interakcji rzędu drugiego, które występują pomiędzy zmiennymi określonymi jako niezależne i zmienną zależną. Aby uniknąć istotnego zmniejszenia stopnia dopasowania,

związanego z usunięciem interakcji pomiędzy zmiennymi niezależnymi, do modelu należy włączyć efekt reprezentujący wszystkie interakcje pomiędzy zmiennymi niezależnymi (Stanisz, 2007)¹. Ostatecznie uwzględniono zależności pomiędzy wydatkami na rekreację i kulturę a dochodem gospodarstwa domowego w przeliczeniu na jednego członka (X_3) oraz liczbą pokoi w mieszkaniu (X_8).

Model log-liniowy przyjmuje wówczas następującą postać:

$$\ln(\hat{n}_{ijklmno}) = \bar{n} + \lambda_i^Y + \lambda_j^{X_1} + \lambda_k^{X_3} + \lambda_l^{X_4} + \lambda_m^{X_5} + \lambda_n^{X_6} + \lambda_o^{X_8} + \lambda_{ik}^{YX_3} + \lambda_{io}^{YX_8} + \lambda_{jklmno}^{X_1X_3X_4X_5X_6X_8} \quad (8)$$

Oszacowany model jest dobrze dopasowany do danych empirycznych. Świadczy o tym wartość statystyki χ^2 największej wiarygodności, która wynosi 172,32 ($p = 0,999$) oraz statystyki χ^2 Pearsona równa 175,22 ($p = 0,999$). Wartości tej statystyki nie są istotne, zatem model jest dobrze dopasowany do danych empirycznych. Potwierdza to także wyk. 1 przedstawiający liczebność obserwowaną względem dopasowanej.

¹ Pominięcie efektu reprezentującego wszystkie interakcje pomiędzy zmiennymi niezależnymi powodowało istotność testów χ^2 Pearsona i χ^2 największej wiarygodności, a zatem konieczność odrzucenia oszacowanego modelu.

Na podjęcie decyzji dotyczącej wydatków na rekreację i kulturę mogą mieć również wpływ interakcje rzędu drugiego następujących par zmiennych charakteryzujących gospodarstwa domowe emerytów (tabl. 3):

- pracujący i dochód w przeliczeniu na jednego członka,
- pracujący i oszczędności,
- pracujący i zadłużenie,
- pracujący i liczba pokoi w mieszkaniu,
- dochód w przeliczeniu na jednego członka i oszczędności,
- dochód w przeliczeniu na jednego członka i zadłużenie,
- dochód w przeliczeniu na jednego członka i sytuacja mieszkaniowa,
- oszczędności i zadłużenie,
- oszczędności i sytuacja mieszkaniowa,
- zadłużenie i liczba pokoi w mieszkaniu,
- sytuacja mieszkaniowa i liczba pokoi.

Czynnikiem, który najrzadziej wchodzi w istotne interakcje z innymi zmiennymi determinującymi podjęcie decyzji o wydatkach na rekreację i kulturę jest sytuacja mieszkaniowa respondenta.

W dalszej części analizy wzięto pod uwagę zmienne opisujące wyposażenie gospodarstw domowych emerytów w sprzęt służący zaspokojeniu potrzeb w zakresie rekreacji i kultury: X_9 , X_{10} , X_{12} , X_{13} , X_{14} , X_{16} . W tym celu dokonano testowania wszystkich interakcji. Okazało się, że w modelu oprócz czynników głównych należy uwzględnić interakcje maksymalnie rzędu drugiego lub trzeciego (tabl. 4). Wskazuje na to prawdopodobieństwo testowe (mniejsze od 0,05) odpowiadające wartości statystyki testowej χ^2 największej wiarygodności i χ^2 Pearsona. Aby sprawdzić, które z zależności powinny być włączone do modelu skorzystano z zależności cząstkowych i brzegowych. Okazało się, że żadna spośród interakcji powyżej rzędu drugiego nie okazała się statystycznie istotna i dlatego w tabl. 5 podano interakcje maksymalnie drugiego rzędu.

TABL. 4. WYNIKI TESTÓW INTERAKCJI MIĘDZY ZMIENNYMI Y , X_9 , X_{10} , X_{12} , X_{13} , X_{14} , X_{16}

Stopień interakcji	Stopnie swobody	Wartość χ^2 największej wiarygodności	prawdopodobieństwo p	Wartość χ^2 Pearsona	prawdopodobieństwo p
1	7	1889,340	0,00000	3389,733	0,00000
2	21	856,042	0,00000	1396,851	0,00000
3	35	49,082	0,05746	53,884	0,02162
4	35	16,698	0,99624	18,783	0,98858
5	21	10,124	0,97726	9,747	0,98202
6	7	4,419	0,73043	4,647	0,70297
7	1	0,235	0,62792	0,235	0,62756

U w a g a. Pogrubioną czcionką zaznaczono istotne interakcje.

Ź r ó d ł o: jak przy tabl. 1.

**TABL. 5. FRAGMENTY WYNIKÓW TESTÓW CZĄSTKOWYCH I BRZEGOWYCH
MIĘDZY ZMIENNYMI $Y, X_9, X_{10}, X_{12}, X_{13}, X_{14}, X_{16}$**

Czynniki	Stopnie swobody	Związek cząstkowy χ^2	Wartość p w związku cząstkowym	Związek brzegowy χ^2	Wartość p w związku brzegowym
Y	1	617,512	0,000	617,512	0,000
X_9	1	46,549	0,000	46,549	0,000
X_{10}	1	358,948	0,000	358,948	0,000
X_{12}	1	4,904	0,027	4,904	0,027
X_{13}	1	776,637	0,000	776,637	0,000
X_{14}	1	76,281	0,000	76,281	0,000
X_{16}	1	8,511	0,004	8,511	0,004
YX_9	1	12,957	0,000	54,574	0,000
YX_{10}	1	8,633	0,003	42,545	0,000
YX_{12}	1	94,819	0,000	124,519	0,000
YX_{13}	1	2,514	0,113	0,621	0,431
YX_{14}	1	0,456	0,499	20,347	0,000
YX_{16}	1	11,117	0,001	11,148	0,001
$X_9 X_{10}$	1	78,975	0,000	173,932	0,000
$X_9 X_{12}$	1	24,001	0,000	79,215	0,000
$X_9 X_{13}$	1	3,251	0,071	26,346	0,000
$X_9 X_{14}$	1	49,042	0,000	132,492	0,000
$X_9 X_{16}$	1	0,427	0,513	14,062	0,000
$X_{10} X_{12}$	1	14,714	0,000	62,224	0,000
$X_{10} X_{13}$	1	7,007	0,008	31,549	0,000
$X_{10} X_{14}$	1	35,412	0,000	115,909	0,000
$X_{10} X_{16}$	1	0,686	0,408	14,227	0,000
$X_{12} X_{13}$	1	6,234	0,013	15,390	0,000
$X_{12} X_{14}$	1	4,015	0,045	25,031	0,000
$X_{12} X_{16}$	1	26,151	0,000	5,767	0,016
$X_{13} X_{14}$	1	21,768	0,000	51,730	0,000
$X_{13} X_{16}$	1	1,155	0,282	8,120	0,004
$X_{14} X_{16}$	1	94,721	0,000	117,107	0,000

U w a g a. Jak przy tabl. 3.

Ż r ó d ł o: jak przy tabl. 1.

Ostatecznie otrzymano następujący model log-liniowy:

$$\ln(\hat{n}_{ijklmn}) = \bar{n} + \lambda_i^Y + \lambda_j^{X_9} + \lambda_k^{X_{10}} + \lambda_l^{X_{12}} + \lambda_m^{X_{13}} + \lambda_n^{X_{14}} + \lambda_n^{X_{16}} + \lambda_{ij}^{YX_9} + \lambda_{im}^{YX_{10}} + \lambda_{in}^{YX_{12}} + \lambda_{in}^{YX_{16}} + \lambda_{jklmn}^{X_9 X_{10} X_{12} X_{13} X_{14} X_{16}} \quad (9)$$

W modelu uwzględniono zależności pomiędzy wydatkami na rekreację i kulturę a posiadaniem:

- aparatu fotograficznego (X_9),
- komputera z dostępem do Internetu (X_{10}),
- urządzenia do odbioru TV satelitarnej lub kablowej (X_{12}),
- roweru (X_{16}).

Oszacowany model jest dobrze dopasowany do danych empirycznych. Świadczą o tym wartość statystyki χ^2 największej wiarygodności, która wynosi 38,911 ($p = 0,968$) oraz wartość statystyki χ^2 Pearsona równa 39,313 ($p = 0,964$). Potwierdza to również układ obserwacji dopasowanych względem obserwacji empirycznych zilustrowany na wyk. 2.

Na podstawie oszacowanego modelu można stwierdzić również występowanie istotnych interakcji rzędu drugiego następujących zmiennych określających posiadanie:

- aparatu fotograficznego i komputera z dostępem do Internetu,
- aparatu fotograficznego i urządzenia do odbioru TV satelitarnej lub kablowej,
- aparatu fotograficznego i samochodu osobowego,
- komputera z dostępem do Internetu i urządzenia do odbioru TV satelitarnej lub kablowej,
- komputera z dostępem do Internetu i działki rekreacyjnej,
- komputera z dostępem do Internetu i samochodu osobowego,
- urządzenia do odbioru TV satelitarnej lub kablowej i działki rekreacyjnej,
- urządzenia do odbioru TV satelitarnej lub kablowej i roweru,
- działki rekreacyjnej i samochodu osobowego,
- samochodu osobowego oraz roweru.

Wynika stąd, że czynnikami, które najczęściej wchodziły w istotne interakcje z innymi zmiennymi opisującymi wydatki na rekreację i kulturę jest posiadanie: samochodu osobowego, komputera z dostępem do Internetu oraz urządzenia do odbioru TV satelitarnej lub kablowej.

Podsumowanie

Zastosowanie modelu log-liniowego pozwoliło na wyodrębnienie czynników, które determinują ponoszenie wydatków na rekreację i kulturę przez gospodarstwa domowe emerytów. Nie wszystkie rozpatrywane w badaniu zmienne — dotyczące zarówno sytuacji ekonomicznej gospodarstw domowych, jak i ich wyposażenia w sprzęt służący zaspokojeniu potrzeb w zakresie rekreacji i kultury — wpływają istotnie na wydatki związane z rekreacją i kulturą. Ponoszenie tych wydatków uzależnione jest przede wszystkim od potencjalnych możliwości finansowych respondentów, a także od ich wyposażenia w niektóre dobra trwałego użytku.

Na podjęcie decyzji dotyczącej wydatków w gospodarstwach domowych emerytów mogą mieć również wpływ interakcje pomiędzy przyjętymi w opracowaniu zmiennymi, niezależnymi czynnikami, które najczęściej wchodziły w istotne interakcje z innymi zmiennymi opisującymi wydatki gospodarstw domowych emerytów. Należą do nich: dochód w przeliczeniu na jednego członka gospodarstwa, zadłużenie gospodarstwa oraz posiadanie: samochodu osobowego, komputera z dostępem do Internetu i urządzenia do odbioru TV satelitarnej lub kablowej. Zmiennymi, które najrzadziej wchodziły w istotne interakcje z pozostałymi czynnikami były sytuacja mieszkaniowa respondenta oraz posiadanie roweru.

Wykorzystanie modelu log-liniowego pozwoliło na bardziej precyzyjny opis zależności między zmiennymi kategorialnymi w porównaniu z miernikami stosowanymi do oceny współzależności cech jakościowych.

dr Iwona Bąk — Zachodniopomorski Uniwersytet Technologiczny w Szczecinie

LITERATURA

- Dobosz M. (2004), *Wspomagana komputerowo statystyczna analiza wyników badań*, Akademicka Oficyna Wydawnicza ELIT, Warszawa
- Salamaga M. (2008), *Wykorzystanie analizy log-liniowej do wyboru czynników opisujących sytuację ekonomiczną gospodarstw domowych*, „Przegląd Statystyczny”, t. 55, nr 4
- Stanisz A. (2007), *Przystępny kurs statystyki z zastosowaniem Statistica PL na przykładach z medycyny*, „Analizy wielowymiarowe”, t. 3, StatSoft, Kraków
- Swedziak K. (2005), *Wykorzystanie bayesowskiego modelu do oceny zmian wzrostu fasoli (Phaseolus Vulgaris)*, Inżynieria Rolnicza, nr 14

- Tomczyk E., Widłak M. (2010), *Konstrukcja i własności hedonicznego indeksu cen mieszkań dla Warszawy*, „Bank i Kredyt”, nr 41 (1)
- Żołnierczuk-Kieliszek D. U., Kulik T. B., Pacian A. B., Stefanowicz A. (2006), *Spółeczno-zdrowotne uwarunkowania decyzji kobiet o stosowaniu hormonalnej zastępczej, Wiadomości Lekarskie*, LIX, nr 9—10

SUMMARY

The aim of article is to select a set of variables that affect decisions about the size of the expenditure incurred on recreation and culture in households of pensioners in Poland. The basis of the information of researches were unidentifiable microdata on income and expenditure of individual households from the household budget survey conducted in 2009 by the GUS (Central Statistical Office) which were associated with the results of a survey "Tourism and recreation in households" conducted by the GUS in the same year. The integrated data set consisted of 1,308 households of pensioners. The study takes into account, first of all, the categorization variables and therefore to a "best" set of factors characterizing household spending on recreation and culture used log-linear analysis. Using log-linear model has enabled more precise description of the relationships between categorization variables compared with indicators used to assess the quality characteristics of interdependence. In addition, log-linear analysis allowed to assess the impact of interactions between variables.

РЕЗЮМЕ

Целью статьи является выбор переменных влияющих на принятие решений в области размера расходов на отдых и культуру в домашних хозяйствах пенсионеров. Основу составляли неидентифицируемые единичные данные ЦСУ по доходам и расходам индивидуальных домашних хозяйств из Обследования бюджетов домашних хозяйств связанные с результатами анкетного обследования «Туризм и отдых в домашних хозяйствах».

В обследовании были приняты во внимание прежде всего категоризационные переменные и поэтому для выбора «оптимального» набора факторов характеризующих расходы домашних хозяйств на отдых и культуру был использован лог-линейный анализ. Это позволило точно описать отношения между категоризационными переменными по сравнению с измерителями используемыми для оценки взаимозависимости качественных особенностей. Лог-линейный анализ позволил также оценить влияние взаимодействий переменных.